COPY No.

OCTOBER 1953
IMM-NYU 200

NEW YORK UNIVERSITY
INSTITUTE FOR MATHEMATICS
AND MECHANICS

Numerical Solution of Flood Prediction and River Regulation Problems

J. J. STOKER

REPORT I

DERIVATION OF BASIC THEORY AND FORMULATION OF NUMERICAL METHODS OF ATTACK

CONTRACT DA-33-017-Eng-223
PREPARED UNDER THE SPONSORSHIP OF
U. S. ARMY ENGINEER CORPS
OHIO RIVER DIVISION

Uncl.

IMM-NYU 200
October 1953

New York University

Institute of Mathematical Sciences

NUMERICAL SOLUTION OF FLOOD PREDICTION AND RIVER REGULATION PROBLEMS

J. J. Stoker

Report I

Derivation of Basic Theory and Formulation of Numerical Methods of Attack

Contract DA-33-017-Eng-223

Prepared under the sponsorship of

U.S. Army Engineer Corps, Ohio River Division

New York, 1953

NUMERICAL SOLUTION OF FLOOD PREDICTION AND RIVER REGULATION PROBLEMS

J. J. Stoker

I. Derivation of Basic Theory and Formulation of Numerical Methods of Attack

1. Introduction

The Institute for Mathematics and Mechanics, New York University, has undertaken the task of investigating the possibility of using modern calculating equipment for solving various problems concerning wave motions in rivers: the problem of predicting floods from the observed and estimated run-off from precipitation, the problem of flood stages at a junction of two large rivers, and the problem of predicting the effects due to operation of a dam are typical examples. This work is being done under a contract with the Corps of Engineers of the U.S. Army in close cooperation with the Ohio River Division in Cincinnati.

The Ohio River Division already has a satisfactory procedure for predicting floods in the Ohio on the basis of the observed, or estimated, run-off and the observed flow into the Ohio from its tributaries; the procedure is called flood-routing. The flood-routing procedure can be deduced as an approximation to the solution of the basic differential equations for flow in open channels (cf. the article by R. R. Gilcrest in the book by Rouse [6](*)), but it makes no

---

(*) Square brackets refer to the bibliography at the end of this report.

direct use of the differential equations. However, the flood-routing procedure in question seems not to give entirely satisfactory results in cases other than that of determining the progress of a flood down a long river--for example, the problem of what happens at a junction, such as that of the Ohio and Mississippi, or the problem of calculating the transient effects resulting from regulation of a dam, such as the Kentucky dam at the mouth of the Tennessee river, seem to be difficult to treat by methods that are modifications of the more or less standard flood-routing procedures. Even for the Ohio, the usual procedure fails occasionally to yield the observed river stages. On the other hand, the basic differential equations for flow in open channels are applicable in all cases and can be used to solve the problems once the appropriate data describing the river and the appropriate initial and boundary conditions are known. The present report has as its objects 1) the derivation of the differential equations and a discussion of a number of properties of their solutions which are of basic importance for a proper understanding of the methods to be used to solve the concrete problems numerically, and 2) a discussion of the numerical methods to be used in solving the problems.

The idea of using the differential equations directly as a means of treating problems of flow in open channels is not at all new. In fact, it goes back to Massau [4] as long ago as 1889. Since then the idea has been taken up by many

others (mostly in ignorance of the work of Massau)--for example, by Preiswerk [5], Thomas [8], Stoker [7]. Thomas, in particular, attacked the flood-routing problem in his noteworthy and pioneering paper [8] and outlined a numerical procedure for its solution based on the idea of using the method of finite differences. However, his method is very laborious to apply and would also not necessarily furnish a good approximation to the desired solution even if a large number of divisions of the river into sections were to be taken. In general, the amount of numerical work to be done looked too formidable for practical purposes until rather recently.

During and since the late war new developments have taken place which make the idea of tackling flood prediction and other similar problems by numerical solution of the relevant differential equations quite tempting. There have been, in fact, developments in two different directions, both motivated by the desire to solve difficult problems in compressible gas dynamics: 1) development of appropriate numerical procedures--for the most part methods using finite differences--for solving the differential equations, and 2) development of computing machines of widely varying characteristics suitable for carrying out the numerical calculations. It happens that the differential equations for flood control problems are of the same type as those for compressible gas dynamics, and consequently the experience and calculating equipment developed for gas dynamics problems

can be used, or suitably modified, for solving flood control problems. The basic general objective of the work of the IMM in this field is to analyze these possibilities and test them by coding concrete problems for suitable calculating machines and comparing the results with observations.

In section 2 a brief derivation of the well-known differential equations governing flow in open channels will be given. In section 3 various interesting and basic general properties of the solutions of the differential equations are discussed on the basis of the theory of characteristics. It is possible to make use of the characteristic form of the differential equations to solve them numerically, but we shall not proceed in such a fashion. Nevertheless an understanding of the theory of the equations based on the use of the characteristic form is vital for the proper understanding and use of the numerical methods we propose to use. In section 4 the numerical methods we employ are described; they are methods in which the field of integration in the x,t-plane is covered by a rectangular net of points, the derivatives in the differential equations are replaced by difference quotients, and the approximate solutions are found by solving linear equations for the values of the desired quantities (essentially the depth of the water and its flow speed) at the net points.

2. Brief derivation of the differential equations

The differential equations which form the basis for our subsequent discussion are well known and many derivations of them have been given (cf., for example, the article by Gilcrest [6]). For the sake of completeness, and to fix notations, we give a brief derivation.

The theory is one-dimensional, i.e. the actual flow in the channel is assumed to be well approximated by a flow with uniform velocity over each cross-section, and the free surface is taken to be a level line in each cross-section. The river is assumed also to be straight enough that its course can be developed into a straight line without causing serious errors in the flow. The flow velocity is denoted by v, the depth of the stream by y, and these quantities are functions of the distance x down the stream and of the time t. (cf. Fig. 2.1). The vertical coordinates of the bottom and of the free surface of the stream, as measured from the horizontal axis x, are denoted by h(x,t) and z(x). The breadth of the free surface at any section of the stream is denoted by B.

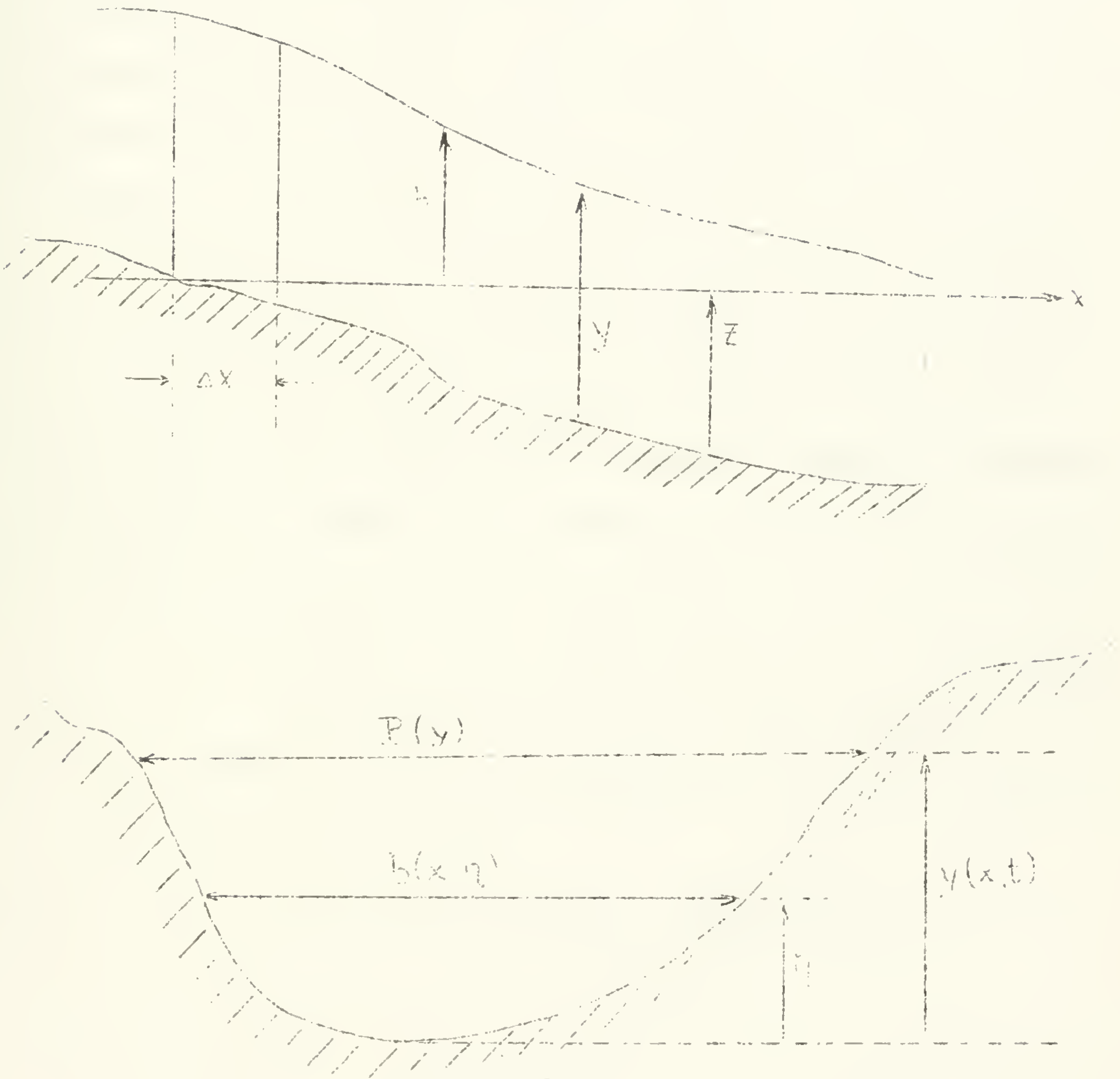


Fig. 2.1

Basic Notations

The differential equations governing the flow are expressions of the laws of conservation of mass and momentum. In deriving them the following assumptions, in addition to those mentioned above, are made(*): 1) the pressure in the water obeys the hydrostatic pressure law, 2) the slope of the bed of the river is small, 3) the effects of friction and turbulence can be accounted for through the introduction of a resistance force depending on the square of the velocity v and also, in a certain way to be specified, on the depth y.

We first derive the so-called equation of continuity from the fact that the mass $\rho A \Delta x$ included in a layer of water of density ρ, thickness Δx, and cross-section area A, changes in its flow along the stream only through a possible inflow along the banks of the stream, say at the rate ρq per unit length along the river: here we use the fact that the particles in a vertical plane section remain in it by virtue of one of the assumptions made above. This means that the so-called particle derivative $\frac{d}{dt}(\rho A \Delta x)$ has the value $\rho q \Delta x$, and hence that the equation

$$(2.1) \qquad v_x \Delta x A + \Delta x (A_t + v A_x) = q \Delta x \quad \text{holds} \quad ,$$

(*) These assumptions are not the minimum number necessary: for example, assumption 1) has as a consequence the independence of the velocity on the vertical coordinate if that were true at any one instant (cf. [7], p. 14 and p. 81).

since $\frac{d\rho}{dt} = 0$, $\frac{d\Delta x}{dt} = v_x \Delta x$ represents the change in thickness of the slice, and the operation $\frac{d}{dt}$ is defined by(*)

(2.2) $$\frac{d}{dt} = \frac{\partial}{\partial t} + v\frac{\partial}{\partial x} \quad ,$$

with v the velocity of the stream. Equation (2.1) can be put in the form

(2.3) $$(Av)_x + A_t = q \quad .$$

Here q represents the volume of the flow into the river over its banks per unit length and time. (This equation could also be derived directly and easily by calculating the flow into and out of the slice of thickness Δx.) For the special case of a rectangular channel of constant breadth B, and hence such that $A = By$, the equation of continuity is at once seen to be expressible as follows:

(2.4) $$v_x y + v y_x + y_t = q/B \quad .$$

The equation of motion is next derived for the same mass included between two vertical planes a distance Δx apart. The momentum of such a mass is given by $\rho A \Delta x v$, and Newton's second law states that

---

(*) If $Q(x,t)$ is any function associated with a particle, i.e. attached to a particle whose motion is given by $x = x(t)$, it follows that $\frac{dQ}{dt} = Q_x \frac{dx}{dt} + Q_t = (\frac{\partial}{\partial t} + v\frac{\partial}{\partial x})Q$, since $\frac{dx}{dt} = v$, the velocity of the particle. Thus, since $\Delta x = x_+ - x_-$, with $x_\pm$ the coordinates of the vertical planes defining our slice, it follows that $\frac{d\Delta x}{dt} = v_+ - v_- \simeq v_x \Delta x$ since x is one of the independent variables. It should be observed that the subscripts x and t are used to denote partial derivatives.

$$\frac{d}{dt}(\rho A \Delta x v) = F \quad , \tag{2.5}$$

with $\frac{d}{dt}$ once more the particle derivative, and F the resultant of all forces on the slice in the horizontal direction. The forces acting on the element in question are indicated in Fig. 2.2. The force P is the resultant pressure on a vertical

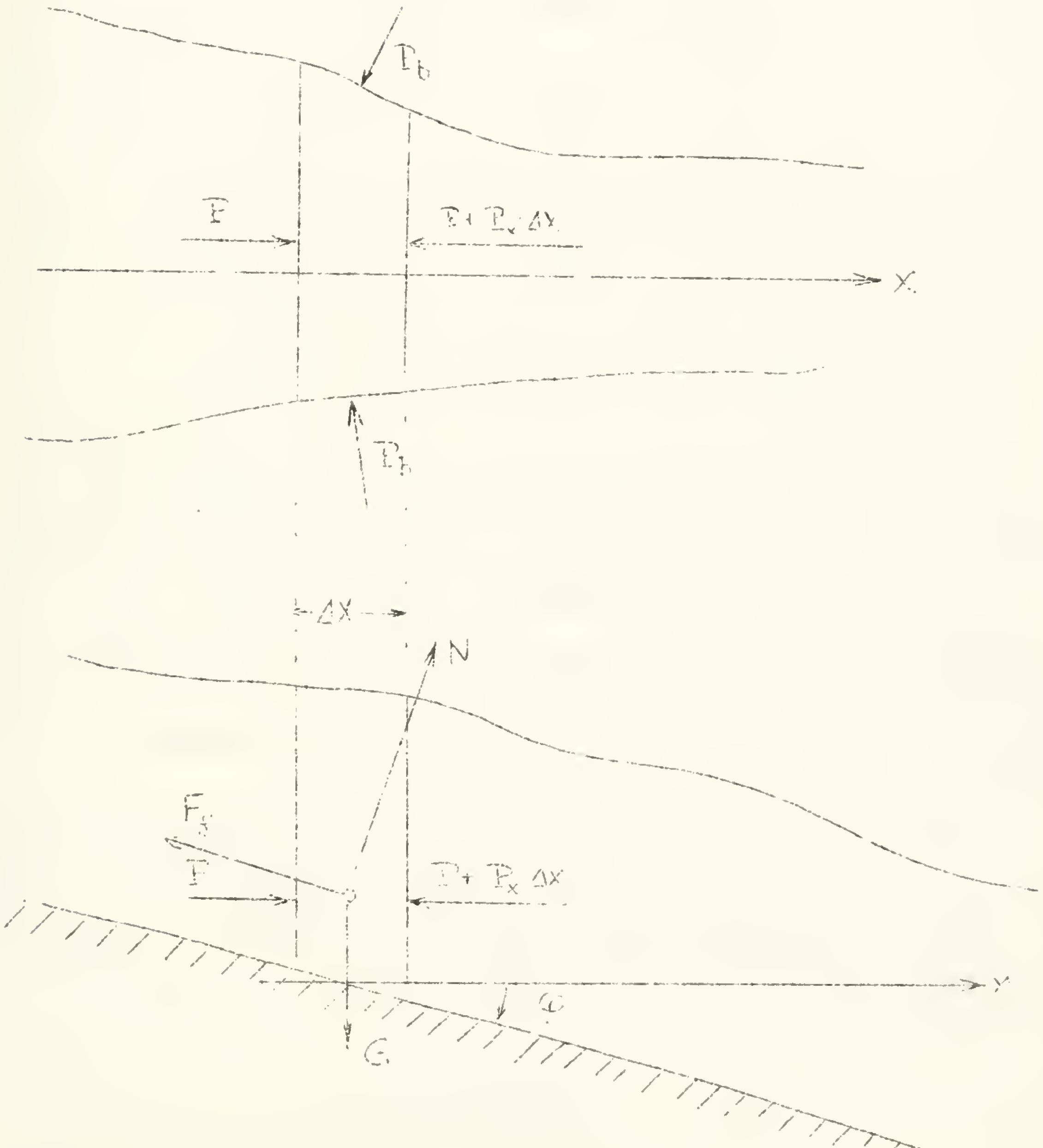


Fig. 2.2

Forces on a Vertical Slice

cross-section, the forces P_b represent the pressure in the horizontal plane from the river banks, G is the weight of the element, F_f the friction force, and N the pressure normal to the river bottom. Because of the hydrostatic pressure assumption we have (cf. Fig. 2.1)

$$(2.6) \qquad P = \int_0^y \rho g(y - \eta) b(x, \eta) d\eta \ ,$$

and hence the quantity P_x is given by

$$(2.7) \qquad P_x = \int_0^y \rho g y_x b(x, \eta) d\eta + \int_0^y \rho g(y - \eta) b_x(x, \eta) d\eta$$

$$= \rho g A y_x + \int_0^y \rho g(y - \eta) b_x(x, \eta) d\eta \quad .$$

The second term on the right can be interpreted as the x-component of the pressure P_b, and is thus balanced by this pressure. Since $\frac{d}{dt}(\rho A \Delta x) = \rho q \Delta x$ is the continuity condition, it follows that we have from (2.5) the equation

$$(2.8) \qquad \rho A \Delta x(v_t + v v_x) + \rho q \Delta x \cdot v = \Delta x N \sin \phi - \Delta x F_f \cos \phi - P_x \Delta x .$$

Since the slope angle ϕ is small, we can replace $\sin \phi$ by $\frac{dz}{dx} = S$, the slope of the river bed, and $\cos \phi$ by 1. In addition, N can be replaced by G--again because ϕ is small --and hence

$$(2.9) \qquad \Delta x N = \Delta x A \rho g \quad .$$

For the resistance force F_f we use Manning's formula:

(2.10) $$F_f = \frac{\rho g A v|v|}{\gamma R^{4/3}} \quad ,$$

which states that the resistance is opposite in direction to the velocity, but is proportional to its square. The constant γ is a roughness coefficient, and R is the hydraulic radius, i.e., the cross-section area divided by the wetted perimeter (cf. [6]). Thus we obtain, finally, the equation of motion:

(2.11) $$\begin{aligned} v_t + vv_x + \frac{q}{A}v &= Sg - gy_x - \frac{1}{\rho A} F_f \\ &= Sg - S_f g - gy_x \quad , \end{aligned}$$

with S_f the so-called friction slope defined by

(2.12) $$S_f = \frac{1}{\rho g A} F_f \quad .$$

The differential equations (2.3) and (2.11) are the differential equations which are basic for the work to follow. Of course, the cross-section area A must be regarded as a known function of y that is fixed by the topography of the river valley; thus (2.3) could also be written in the form

(2.13) $$A_y y_x v + A_x v + Av_x + A_y y_t = q \quad ,$$

and (2.11) and (2.13) then are a pair of equations to determine the two quantities $y(x,t)$ and $v(x,t)$.

For any given river it is thus necessary to have data available capable of yielding the cross-section area A, the hydraulic radius R, and the roughness coefficient γ of the river

as functions of x and y, and of the slope S of its bed as a function of x in order to have the coefficients in the differential equations (2.11) and (2.13) defined. Three of these quantities are purely geometrical in character and could in principle be determined from an accurate contour map of the river valley, but the determination of the roughness coefficient γ of course requires measurements of actual flows in the river for its determination. (These remarks imply that the river does not change its course from time to time--as happens, for example, in the Missouri--so that the coefficients of the differential equations are independent of the time.)

3. The method of characteristics

It has already been stated that many important properties of the solutions of our differential equations and many valuable insights with regard to the integration theory associated with them can be gained through the use of what is called the method of characteristics. In this section we give an explanation of this method, but since we do not make direct use of it later on in solving problems numerically we shall base our explanation on a special case of the differential equations, although the results and observations are applicable in the most general case. The special case in question is that of a river of constant rectangular section and uniform slope, with no flow into the river from the banks (i. . q = 0 in (2.11) and (2.13)). In this case the differential

equations (2.11) and (2.13), with $A = By$, can be written as follows:

$$(3.1) \qquad v_x y + v y_x + y_t = 0 \quad ,$$

$$(3.2) \qquad v_t + v v_x + g y_x + E = 0 \quad .$$

We have introduced the symbol E for the external forces per unit mass:

$$(3.3) \qquad E = -gS + gS_f \quad , \quad S = \text{const.}$$

The term E differs from the others in that it contains no derivatives of y or v.

The theory of characteristics for these equations can be approached very directly(*) in the present special case by introducing a new quantity c to replace y, as follows:

$$(3.4) \qquad c = \sqrt{gy} \quad .$$

This quantity has great physical significance, since it represents--as we shall see--the propagation speed of small disturbances in the river. From (3.4) we obtain at once the relations

$$(3.5) \qquad 2cc_x = gy_x \quad , \quad 2cc_t = gy_t \quad ,$$

and the differential equations (3.1) and (3.2) take the form

(*) For a treatment which shows quite generally how to arrive at the formulation of the characteristic equations, see [1], Ch. 2.

$$(3.6)\qquad \begin{cases} 2cc_x + v_t + vv_x + E = 0 \quad , \\ cv_x + 2vc_x + 2c_t = 0 \quad . \end{cases}$$

These equations are next added, then subtracted, to obtain the following equivalent pair of equations:

$$(3.7)\qquad \begin{cases} 2\left\{(c+v)\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\right\} c + \left\{(c+v)\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\right\} v + E = 0 \quad , \\ -2\left\{(-c+v)\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\right\} c + \left\{(-c+v)\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\right\} v + E = 0 \quad . \end{cases}$$

We observe that the derivatives in these equations now have the form of directional derivatives--indeed, to achieve that was the purpose of the transformation--so that c and v in the first equation, for example, are both subject to the operator $(c+v)\frac{\partial}{\partial x} + \frac{\partial}{\partial t}$, which means (cf. the footnote in the preceding section) that these functions are differentiated along curves in the x,t-plane which satisfy the differential equation $\frac{dx}{dt} = c+v$. In similar fashion, the functions c and v in the second equation are both subject to differentiation along curves satisfying the differential equation $\frac{dx}{dt} = -c+v$.

It is entirely feasible to develop the integration theory of equations (3.7) quite generally on the basis of these observations (as is done, for example, in [1], Ch. 2), but it is simpler, and leads to the same general results, to describe it for the special case in which the resistance force F_f is neglected so that the quantity E in (3.7) is a constant (see (3.3)). In this case the equations (3.7) can be written in the form

$$(3.7)_1 \quad \begin{aligned} \left\{\frac{\partial}{\partial t} + (v+c)\frac{\partial}{\partial x}\right\}(v+2c+Et) &= 0 \quad , \\ \left\{\frac{\partial}{\partial t} + (v-c)\frac{\partial}{\partial x}\right\}(v-2c+Et) &= 0 \quad , \end{aligned}$$

as one can readily verify. But the interpretation of the operations defined in $(3.7)_1$ has just been discussed: The relations state that the functions $(v \pm 2c + Et)$ are constant for a point moving through the fluid with the velocity $(v \pm c)$, or, as we may also put it, for a point whose motion in the x,t-plane is characterized by the ordinary differential equations $\frac{dx}{dt} = v \pm c$. That is, we have the following situation in the x,t-plane: There are two sets of curves, C_1 and C_2 called characteristics, which are the solution curves of the ordinary differential equations

$$(3.8) \quad \begin{cases} C_1: \quad \dfrac{dx}{dt} = v+c \quad , \quad \text{and} \\ C_2: \quad \dfrac{dx}{dt} = v-c \quad , \end{cases}$$

and we have the relations

$$(3.9) \quad \begin{cases} v+2c+Et = k_1 = \text{const. along a curve } C_1 \text{ and} \\ v-2c+Et = k_2 = \text{const. along a curve } C_2 \; . \end{cases}$$

Of course the constants k_1 and k_2 will be different on different curves in general. It should also be observed that the two families of characteristics determined by (3.8) are really distinct because of the fact that $c = \sqrt{gy} \neq 0$ since we suppose that $y > 0$, i.e. that the water surface never touches the bottom.

By reversing the above procedure it can be seen rather easily that the system of relations (3.8) and (3.9) is completely equivalent to the system of equations (3.6) for the case of constant bottom slope and zero resistance, so that a solution of either system yields a solution of the other. In fact, if we set $f(x,t) = v+2c+Et$ and observe that $f(x,t) = k_1 = \text{const.}$ along any curve $x = x(t)$ for which $\frac{dx}{dt} = v+c$ it follows that along such curves

$$(3.10) \qquad f_t + f_x \frac{dx}{dt} = f_t + (v+c)f_x = 0 \quad .$$

In the same way the function $g(x,t) = v-2c+Et$ satisfies the relation

$$(3.11) \qquad g_t + (v-c)g_x = 0$$

along the curves for which $\frac{dx}{dt} = v-c$. Thus wherever the curve families C_1 and C_2 cover the x,t-plane the relations (3.10) and (3.11) hold. If now equations (3.10) and (3.11) are added and the definitions of $f(x,t)$ and $g(x,t)$ are recalled it is readily seen that the first of equations (3.6) results. By subtracting (3.11) from (3.10) the second of equations (3.6) is obtained. In other words, any functions v and c which satisfy the relations (3.8) and (3.9) will also satisfy (3.6) and the two systems of equations are therefore now seen to be completely equivalent.

As we would expect on physical grounds, a solution of the original dynamical equations (3.6) could be shown to be uniquely determined when appropriate initial conditions

(for $t = 0$, say) and boundary conditions are prescribed; it follows that any solutions of (3.8) and (3.9) are also unique when such conditions are prescribed since we know that the two systems of equations are equivalent.

At first sight one might be inclined to regard the relations (3.8) and (3.9) as more complicated than the original pair of partial differential equations, particularly since the right hand sides of (3.8) are not known and hence the characteristic curves are also not known. Nevertheless, the formulation in terms of the characteristics is quite useful in studying properties of the solutions and also in studying questions referring to the appropriateness of various boundary and initial conditions. It is useful to begin by describing briefly a method of determining the characteristics by successive approximations which at the same time makes possible a number of useful observations and interpretations regarding the role of the characteristics in general. Let us consider a problem in which the values of the velocity v and the surface elevation y (or, what amounts to the same thing, the wave speed $c = \sqrt{gy}$) are prescribed for all values of x when $t = 0$. We wish to calculate the solution for $t > 0$ by determining v and c through use of (3.8) and (3.9) and the given initial conditions. We have for $t = 0$ the conditions

$$(3.12) \qquad \begin{cases} v(x,0) = \bar{v}(x) \\ c(x,0) = \bar{c}(x) \end{cases}$$

in which $\bar{v}(x)$ and $\bar{c}(x)$ are given functions. We can approximate the values of v and c for small values of t as follows: consider a series of points on the x-axis (cf. Fig. 3.1) a small distance Δx apart. At all of these points the values of v and c are known from (3.12). Consequently the slopes of the characteristics C_1 and C_2 at these points are known from (3.8). From the points 1, 2, 3. 4 straight line segments with these slopes are drawn until they intersect at points 5, 6, and 7, and if Δx is chosen sufficiently small it is reasonable to expect that the positions of these points will be good approximations

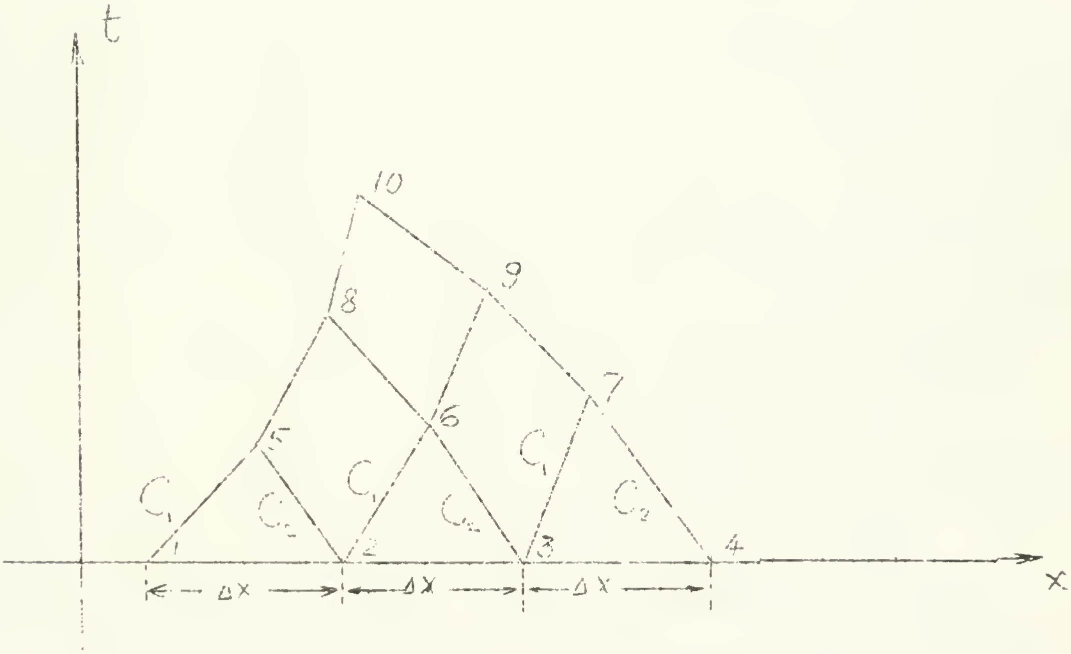


Fig. 3.1

to the intersections of the characteristics issuing from the points 1, 2, 3, 4. The values of both x and t at points 5, 6, and 7 are now known--they can be determined graphically for example--and through the use of (3.9) and the initial

conditions we can also determine the approximate values of v and c at these points. For this purpose we observe that along any particular segment issuing from the points 1, 2, 3 or 4 the values of v+2c + Et and v-2c + Et are known constants since the values of v and c are fixed by (3.12) for t = 0; hence we have

$$(3.13) \qquad \begin{cases} \text{Along } C_1: \; v+2c+Et = \bar{v} + 2\bar{c} \quad , \quad \text{and} \\ \\ \text{Along } C_2: \; v-2c+Et = \bar{v} - 2\bar{c} \quad . \end{cases}$$

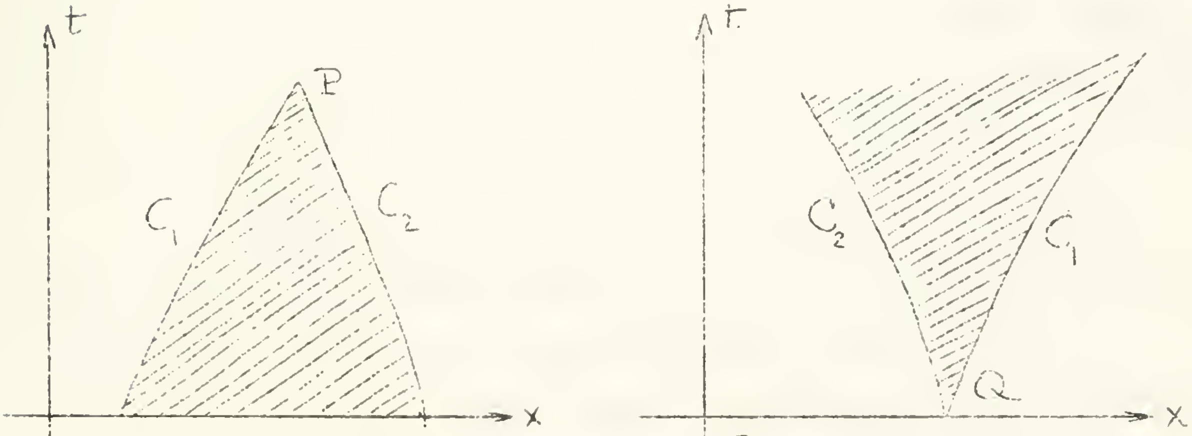
At the points 5, 6, and 7 we know the values of t and hence (3.13) furnishes two independent linear equations for the determination of the values of v and c at each of these points. Once v and c are known at points 5, 6, and 7 the slopes of the characteristics issuing from these points can be determined from (3.8) and the entire process can be carried out once more to yield the additional points 8 and 9 and the approximate values of v and c at these points. In this way we can approximate the values of v and c at the points of a net over a certain region of the x,t-plane, and can then obtain approximate values for v and c at any points in the same region either by interpolation or by refining the net inside the region. It is quite plausible and could be proved mathematically that the above process would converge as $\Delta x \rightarrow 0$ to the unique solution of (3.8) and (3.9), and hence also of (3.6), corresponding to the given initial conditions for sufficiently small values of t (i.e. for a region of the

x,t-plane not too far from the x-axis) provided that the prescribed initial values of v and c are sufficiently regular functions of x.

It should be clear that once the characteristics are known the values of v and c for all points of the x,t-plane covered by them are also known, since the constants k_1 and k_2 in (3.9) are known on each characteristic through the initial data and hence the values of v and c for any point (x,t) can be calculated by solving the linear equations (3.9) for the characteristics through that point. This statement of course implies that each one of the two families of characteristics covers a region of the x,t-plane simply and that no two members of different families are tangent to each other--in other words it is implied that the two families of characteristics form a non-singular curvilinear coordinate system over the region of the x,t-plane in question.

The method of finite differences used above to determine the characteristics can be interpreted in such a way as to throw light on the physical properties of the solution. Consider the point 10 of Fig. 3.1 for example. We recall that the approximate values v_{10} and c_{10} of v and c at point 10 were obtained through making use of the initial values of v and c at points 1, 2, 3, 4 on the x-axis only, and furthermore that the values v_{10} and c_{10} required the use of points confined solely to the region within the approximate characteristics joining point 10 with points 1 and 4. Since the finite difference scheme outlined above converges as $\Delta x \longrightarrow 0$ to

yield the exact characteristics it is clear that we can make the following important statement: the values of v and c at any point P(x,t) within the region of existence of the solution are determined solely by the initial values prescribed on the segment of the x-axis which is subtended by the two characteristics issuing from P. In addition, the two characteristics issuing from P are also determined solely by the initial values on the segment subtended by them. Such a segment of the x-axis is often called the domain of dependence of the point P. Correspondingly we may define the range of influence of a point Q on the x-axis as the region of the x,t-plane in which the values of v and c are influenced by the initial values assigned to point Q. In Fig. 3.2 we indicate these two regions.



Domain of Dependence

Range of Influence

Fig. 3.2

We are now in a position to understand the role of the characteristics as curves along which discontinuities in the first and higher derivatives of the initial data are propagated, since it is reasonable to expect (and could be proved) that those points P whose domains of dependence do not contain such discontinuities are points at which the solutions v and c also have continuous derivatives. On the other hand, it could be shown that a discontinuity in the initial data at a certain point does not in general die out along the characteristic issuing from that point. Such a discontinuity (or disturbance in the water) therefore spreads in both directions over the surface of the water with the speed v+c in one direction and v-c in the other in view of the interpretation given to the characteristics through the relations $(3.7)_1$. Since v is the velocity of the water particles we see that c represents quite generally the speed at which a discontinuity in a derivative of the initial data propagates relative to the moving water. We are therefore justified in referring to the quantity $c = \sqrt{gy}$ as the wave speed or propagation speed.

We considered above a problem in which only initial conditions, and no boundary conditions, were prescribed. In the problems we consider later, however, such boundary conditions will occur in the form of conditions prescribed at a certain fixed point of the river in terms of the time: for example, the height, or stage, of the river might be given at a certain station as a function of the time. In other words, conditions would be prescribed not only along the x-axis of

our x,t-plane, but also along the t-axis (in general only for t > 0) for a certain fixed value of x. The method of finite differences used above to discuss the initial value problem can be modified in an obvious way to deal with cases in which boundary conditions are also imposed. In doing so, it would also become clear just what kind of boundary conditions could and should be imposed. For example, in the great majority of rivers--in fact, for all in which the flow is sub-critical, i.e. such that v is everywhere less than the wave speed $\sqrt{gy}$ --it is possible to prescribe only one condition along the t-axis, which might be either the velocity v or the depth y, in contrast with the necessity to impose two conditions along the x-axis. This fact would become obvious on setting up the finite difference scheme, and examples of it will be seen in later reports.

Finally, it should be stated that the role of the characteristics, and also the method of finite differences applied to them could be used with reference to the general case of the characteristic equations as embodied in the equations (3.7) and (3.8) in essentially the same way as was sketched out above for the system comprised of (3.8) and (3.9) which referred to a special case. In particular, the role of the characteristics as curves along which small disturbances propagate, and their role in determining the domain of dependence, range of influence, etc. remain the same.

4. Approximate methods to be used in calculating solutions of the differential equations

It has already been stated that while the formulation of our problems by the method of characteristics is most valuable for studying many questions concerned with the solutions of the differential equations, it is in most cases not the best formulation to use for the purpose of calculating the solutions numerically. That is not to say that the device of replacing derivatives by difference quotients is given up, but rather that this device is used in a different manner. The basic idea is to operate with finite differences by using a fixed rectangular net in the x,t-plane, in contrast with the method outlined above, in which the net of points in the x,t-plane at which the solution is to be approximated is determined only gradually in the course of the computation. In the latter procedure it is thus necessary to calculate not only the values of the unknown functions v and c, but also the values of the coordinates x,t of the net points themselves, whereas a procedure making use of a fixed net would require the calculation of v and c only, and it would also have the advantage of furnishing these values at a convenient set of points.

However, the question of the convergence of the approximate solution to the exact solution when the mesh width of a rectangular net is made to approach zero is more delicate than it is when the method of characteristics is used. For example, it is not correct, in general, to choose a net in

which the ratio of the mesh width Δt along the t-axis and the mesh width Δx along the x-axis is kept constant: such a procedure would not in general yield approximations converging to the solution of the differential equation problem. The reason for this can be understood with reference to one of the basic facts about the solutions of the differential equations which was brought out in the discussion of the preceding section. The basic fact in question is the existence of what was called there the domain of dependence of the solution. For example, suppose the solution were to be approximated at the points of the net of Fig. 4.1a by advancing from one row parallel to the x-axis to the next row a distance Δt from it. In addition, suppose this were to be done by determining the approximate values of v and c at any point such as P (cf. Fig. 4.1b) by using the values

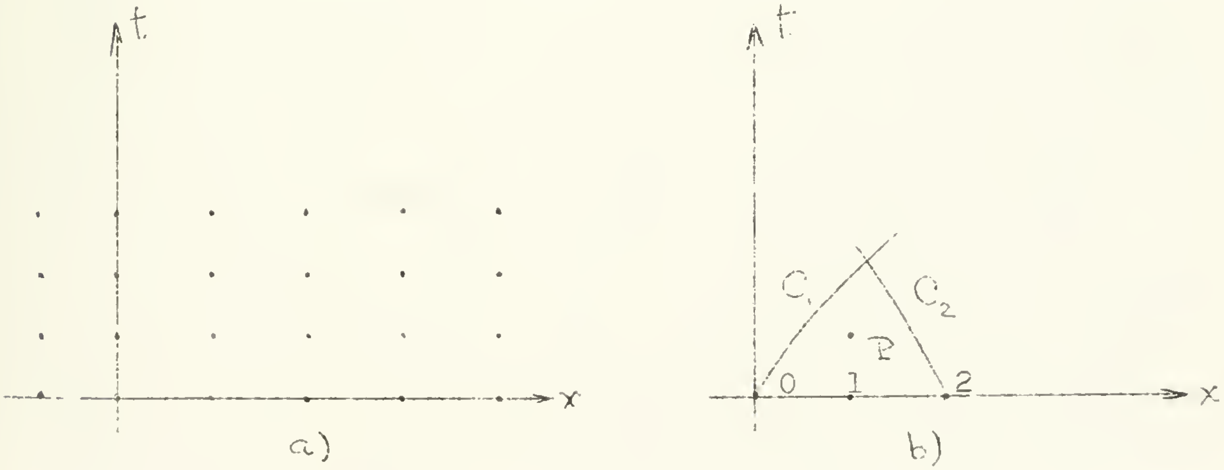


Fig. 4.1

of these quantities at the nearest three points 0, 1, 2 in the next line below, replacing derivatives in the two differential equations by difference quotients, and then solving the resulting algebraic equations for the two unknowns v_P and c_P. It seems reasonable to suppose that such a scheme would be appropriate only if P were in the triangular region formed by the characteristics drawn from points 0 and 2 to form the region within which the solution is determined solely by the data given on the segment 0-2: otherwise it seems clear that the values at additional points on the x-axis ought to be utilized since our basic theory tells us that the initial data at some of them would indeed influence the solution. On the other hand, the characteristic curves depend themselves upon the values of the unknown functions v and c --their slopes, in fact, are given (cf. (3.3)) by $\frac{dx}{dt} = v \pm c$ and thus the interval Δt must be chosen small enough in relation to a fixed choice of the interval Δx so that the points such as P will fall within the appropriate domains of determinacy relative to the points used in calculating the solution at P. In other words, the theory of the characteristics, even if it is not used directly, comes into play in deciding the relative values of Δt and Δx for convergence. (This, by the way, is the reason for the statement made earlier that the finite difference scheme proposed by Thomas [8] might fail to converge no matter how small the mesh width might be taken, since he ignores the facts pointed out here.)

We shall introduce two different schemes employing the method of finite differences in a fixed rectangular net of the x,t-plane. The first of these makes use of the differential equations in the form given by (3.7), and we no longer suppose that the function E is restricted in any way. (It might be noted that the slopes of the characteristics as given by (3.8) are determined by the quantities v ± c, no matter how the function E is defined, and hence we are in a position to determine appropriate lengths for the t-intervals, in accord with the above discussion, in the most general case. This is also a good reason for working with the quantity c in place of y.) At the same time, the calculation is based on assuming that the approximate values of c and v have been calculated at the net points L, M, R (cf. Fig. 4.2) and that the differential equations are to be used to advance the

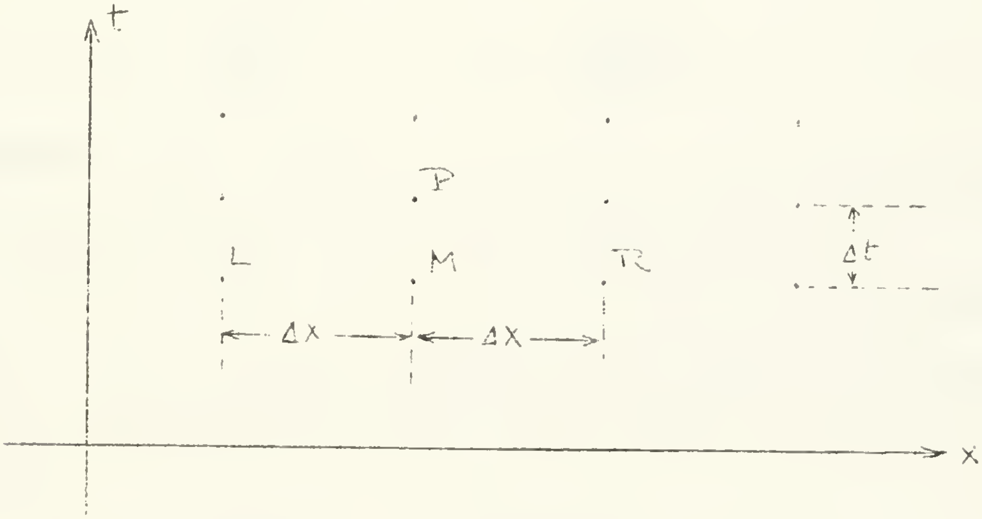


Fig. 4.2

A Rectangular Net

approximate solution to the point P. The differential equations to be solved are thus

$$(4.1) \qquad 2\{(c+v)c_x + c_t\} + \{(c+v)v_x + v_t\} + E = 0 \quad ,$$

$$(4.2) \qquad -2\{(-c+v)c_x + c_t\} + \{(-c+v)v_x + v_t\} + E = 0 \quad ,$$

and the characteristic directions are determined by $\frac{dx}{dt} = v \pm c$. The characteristic with slope v+c we call the forward characteristic, and that with slope v-c the backward characteristic. We shall replace the derivatives in the equations by difference quotients which approximate the values of the derivatives at the point M. In order to advance the values of v and c from the points L, M, R to the point P by using (4.1) and (4.2) it is natural to replace the time derivatives v_t and c_t by the following difference quotients:

$$(4.3) \qquad v_t = \frac{v_P - v_M}{\Delta t} \quad , \qquad c_t = \frac{c_P - c_M}{\Delta t}$$

in both equations. However, in order to insure the convergence of the approximations to the exact solution when $\Delta x \to 0$ and $\Delta t \to 0$ (see [2] for a proof of this fact) it is necessary to replace the derivatives v_x and c_x by difference quotients which are defined differently for (4.1) than for (4.2), as follows:

$$(4.4) \qquad v_x = \frac{v_M - v_L}{\Delta x} \quad , \quad c_x = \frac{c_M - c_L}{\Delta x} \quad \text{in } (4.1) \quad ,$$

$$(4.5) \qquad v_x = \frac{v_R - v_M}{\Delta x} \quad , \quad c_x = \frac{c_R - c_M}{\Delta x} \quad \text{in } (4.2) \quad .$$

The reason for this procedure is, at bottom, that (4.1) is an equation associated with the forward characteristic, while (4.2) is associated with the backward characteristic. The coefficients of the derivatives and the function E are, of course, evaluated at the point M. The difference equations replacing (4.1) and (4.2) are thus given by

$$(4.6)\qquad 2\left\{(c_M+v_M)\frac{c_M-c_L}{\Delta x}+\frac{c_P-c_M}{\Delta t}\right\} + \left\{(c_M+v_M)\frac{v_M-v_L}{\Delta x}+\frac{v_P-v_M}{\Delta t}\right\} + E(v_M,c_M) = 0 \quad ,$$

$$(4.7)\qquad -2\left\{(-c_M+v_M)\frac{c_R-c_M}{\Delta x}+\frac{c_P-c_M}{\Delta t}\right\} + \left\{(-c_M+v_M)\frac{v_R-v_M}{\Delta x}+\frac{v_P-v_M}{\Delta t}\right\} + E(v_M,c_M) = 0 \quad .$$

We observe that the two unknowns, v_P and c_P, occur linearly in these equations; hence they are easily found by solving the equations. The result is

$$(4.8)\qquad v_P = v_M + \frac{\Delta t}{\Delta x}\Big[(c_M+v_M)\left(\tfrac{1}{2}\,v_L - \tfrac{1}{2}\,v_M + c_L - c_M\right) - (c_M-v_M)\left(\tfrac{1}{2}\,v_M - \tfrac{1}{2}\,v_R - c_M + c_R\right) - \Delta x E_M\Big] \quad ,$$

$$(4.9)\qquad c_P = c_M + \frac{1}{2}\,\frac{\Delta t}{\Delta x}\Big[(c_M+v_M)\left(\tfrac{1}{2}\,v_L - \tfrac{1}{2}\,v_M + c_L - c_M\right) + (c_M-v_M)\left(\tfrac{1}{2}\,v_M - \tfrac{1}{2}\,v_R - c_M + c_R\right)\Big] \quad .$$

In accordance with the remarks made above, we must also require that the ratio of Δt to Δx be taken small enough so that P lies within the triangle formed by drawing lines from L and R in the directions of the forward and backward characteristics respectively, i.e. lines with the slopes v_L+c_L at L and v_R-c_R at R: A condition that is well-determined since the values of v and c are presumably known at L and R.

One can now see in general terms how the initial value problem starting at t = 0 can be solved approximately: One starts with a net along the x-axis with spacing Δx. Since c and v are known at all of these points, the values of c and v can be advanced through use of (4.6) and (4.7) to a parallel row of points along a line distant Δt along the t-axis from the x-axis. However, the mesh width Δt must be chosen small enough so that the convergence condition is satisfied at all net points where new values of v and c are computed.

We can now see also how to take care of boundary conditions, i.e. of conditions imposed at a fixed point (say at the origin, x = 0) as given functions of the time. For example, the depth y (corresponding to the stage of the river) or the velocity v (which together with the cross-section area A fixes the rate of discharge) might be given in terms of the time. Initial conditions downstream from this point (i.e. for x > 0) might also be prescribed. Suppose, for example, that the stage of the river is prescribed at x = 0, i.e. that y(0,t)

is known, and that the calculation had already progressed so far as to yield values of v and c at net points along a certain line parallel to the x-axis and containing the points L, M, R, as indicated in Fig. 4.3. It is clear that the

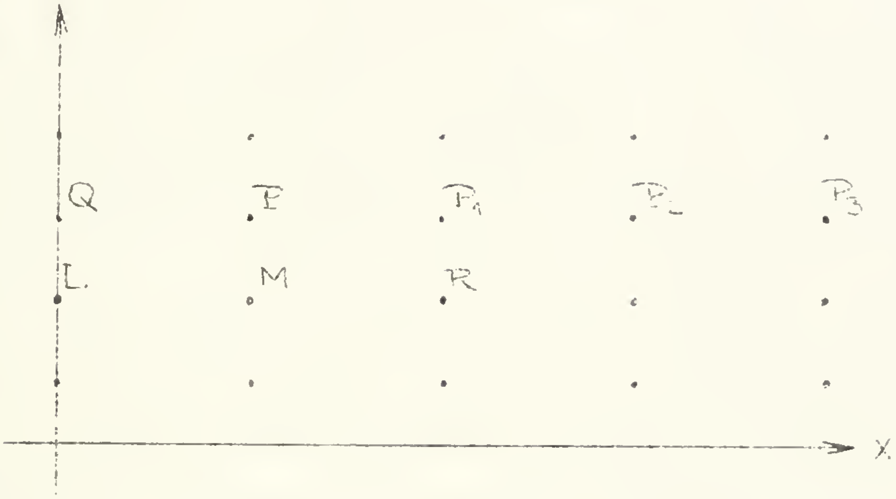


Fig. 4.3

Satisfying Boundary Conditions

determination of the values of v and c at point P can be obtained from their values at L, M, R by using (4.8) and (4.9), as in the above discussion of the initial value problem, and similarly for points P_1, P_2, etc. However, the value of v at Q must be determined in a different manner; for this purpose we use the equation (4.7) with the subscript Q replacing P, L replacing M, and M replacing R. Since v_M, c_M, v_R, c_R are supposed known, and c_Q is also known since the values of y are prescribed on the t-axis, it follows that equation (4.7) contains v_Q as the only unknown; in fact it is given by the equation

$$(4.10)\quad v_Q = v_L + \Delta t\left\{-\frac{1}{\Delta x}(c_L - v_L)(2c_L - 2c_M - v_L + v_M) - E_L\right\} + 2(c_Q - c_L) \ .$$

The reason for using (4.7) instead of (4.6) is, of course, that the points M and Q are associated with the backward characteristic, and hence (4.2) should be used to approximate the x-derivatives at point L (where the differential equations are replaced by difference equations). It is quite clear that the same general procedure could be used to calculate c_Q if the values of v had been assumed given along the t-axis. If, on the other hand, we had a boundary condition on the right instead of on the left, as above, we could make use of (4.6) for the forward characteristic as a basis for obtaining the formula for advancing the solution along the t-axis.

The above discussion would seem to imply that under all circumstances only one boundary condition could be imposed --that is, that either v or c could be prescribed at a fixed point on the river, but not both--since prescribing one of these quantities leads to a unique determination of the other. This is, indeed, true in any ordinary river, but not necessarily in all cases. In fact, we made a tacit assumption in the above discussion, i.e. that of the two characteristics issuing from any point of the t-axis only the forward characteristic goes into the region x > 0 to the right of the t-axis, and this in turn implies that v+c and v-c, which fix the slopes of the characteristics, are opposite in sign. The physical interpretation of this is that the value of v (which is positive here) must be less than $c = \sqrt{gy}$, i.e. that the flow

must be what is called tranquil, or sub-critical.(*) Otherwise, as we see from Fig. 4.4, we should expect to determine the values of v and c at points close to and to the right of the t-axis, say at K, by utilizing values for both v and c along the segment LQ, its domain of dependence. The scheme outlined above would therefore have to be modified under such circumstances. Later on, we may wish to treat flows in streams so precipitous that super-critical velocities arise (as they do in the upper reaches of the New River in West Virginia, and the Los Angeles River even in its lower section), and appropriate finite difference schemes for such cases will then be introduced. One sees, however, how useful the theory based on the characteristics can be even though no direct use

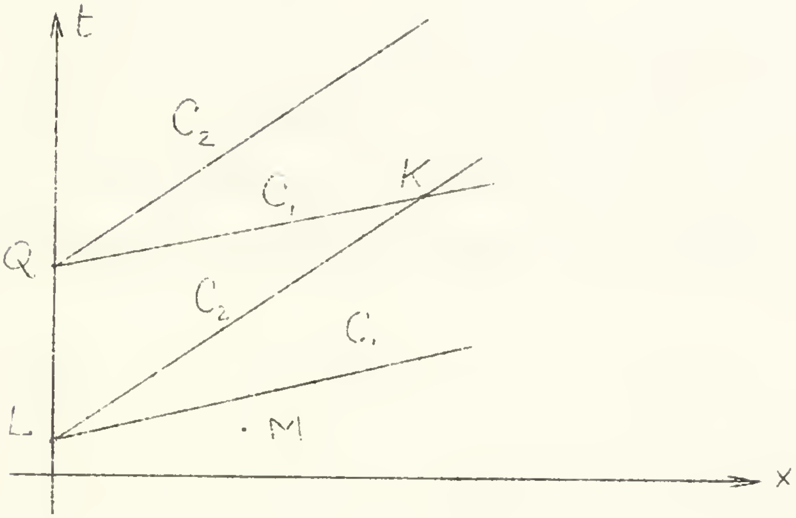


Fig. 4.4

A Case of Super-critical Flow

(*) In gas dynamics the flow in an analogous case would be called sub-sonic.

of it is made in the numerical calculations (aside from decisions on the size of the t-interval).

The procedure sketched out above, while it is recommended for use when a boundary condition is to be satisfied, is not the best one to use for advancing the solution to such points as $P, P_1, P_2, \ldots$ in Fig. 4.3. For such "interior points" a staggered rectangular net, as indicated in Fig. 4.5, and a difference equation scheme based on the original differential

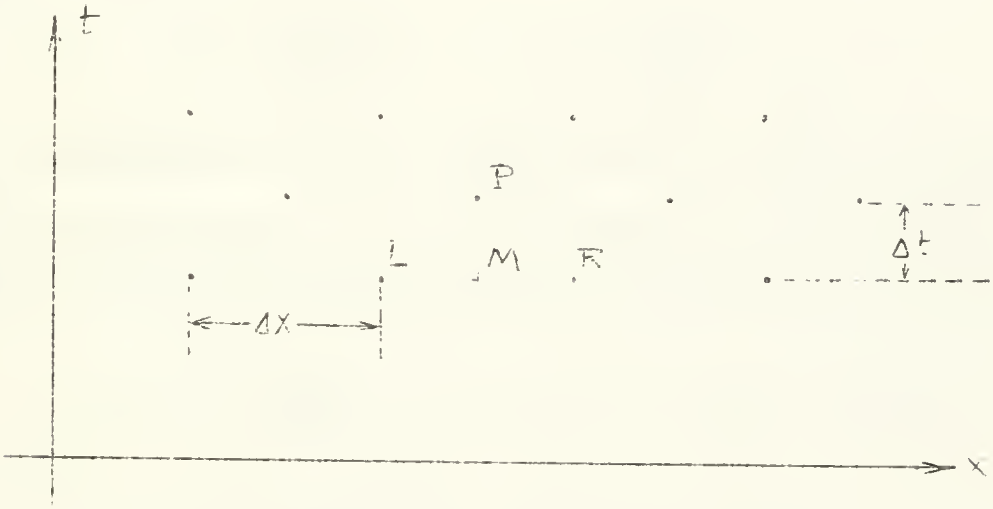


Fig. 4.5

A Staggered Net

equations (3.6) is preferable (cf. [3] for a discussion of this scheme). The equations (3.6) were

$$\tag{4.11} \begin{cases} 2cc_x + v_t + vv_x + E = 0 \quad , \\ cv_x + 2vc_x + 2c_t = 0 \quad . \end{cases}$$

The values v_M and c_M at the mid-point M (which is, however, not a net point) of the segment LR are defined by the averages:

$$v_M = \frac{v_L + v_R}{2} \quad , \quad c_M = \frac{c_L + c_R}{2} \quad , \tag{4.12}$$

after which the derivatives at M are approximated in a quite natural way, by the difference quotients

$$\begin{cases} v_x = \dfrac{v_R - v_L}{\Delta x} \quad , & c_x = \dfrac{c_R - c_L}{\Delta x} \quad , \\ v_t = \dfrac{v_P - v_M}{\Delta t} \quad , & c_t = \dfrac{c_P - c_M}{\Delta t} \quad . \end{cases} \tag{4.13}$$

Upon substitution of these quantities into (4.11), evaluation of the coefficients c, v, and E at point M, and subsequent solution of the two equations for v_P and c_P the result is

$$\begin{cases} v_P = v_M + \dfrac{\Delta t}{\Delta x}[2(c_L - c_R)c_M + (v_L - v_R)v_M - \Delta x E_M] \quad , \\ c_P = c_M + \dfrac{1}{2}\dfrac{\Delta t}{\Delta x}[2(c_L - c_R)v_M + (v_L - v_R)c_M] \quad . \end{cases} \tag{4.14}$$

As we see on comparison with (4.8) and (4.9), these equations are simpler than the earlier ones. The criterion for convergence remains the same as before, i.e. that P should lie within a triangle formed by the segment LR and the two characteristics issuing from its ends.

Bibliography

[1] R. Courant and K. O. Friedrichs, Supersonic Flow and Shock Waves, Interscience, New York, 1948.

[2] R. Courant, E. Isaacson and M. Rees, On the Solution of Nonlinear Hyperbolic Differential Equations by Finite Differences, Communications on Pure and Applied Mathematics, Vol. V, p. 243, 1952.

[3] J. Keller and P. Lax, Finite Difference Schemes for Hyperbolic Systems, LAMS 1201, 1950.

[4] J. Massau, Mémoire sur l'intégration graphique des équations aux dérivées partielles, Annales des ingénieurs sortis de Gand, Tome XII, 1889.

[5] E. Preiswerk, Anwendung gasdynamischer Methoden auf Wasserströmungen mit freier Oberfläche, Mitteilungen aus dem Institut für Aerodynamik, Technische Hochschule, Zürich, 1938.

[6] H. Rouse, Engineering Hydraulics, New York and London, 1950. (Especially the article by B. R. Gilcrest, p. 635).

[7] J. J. Stoker, The Formation of Breakers and Bores, Communications on Applied Mathematics, Vol. I, p. 1, 1948.

[8] H. A. Thomas, The Hydraulics of Flood-Movements in Rivers, Carnegie Institute of Technology, Pittsburgh, Pa., 1937.